

Oponentský posudek na diplomovou práci:

## Kateřina Koňasová: Stochastická rekonstrukce bodových vzorků

---

Diplomová práce se zabývá stochastickou rekonstrukcí bodových vzorků, neboli konfigurací konečného počtu bodů v kompaktním okně v  $\mathbb{R}^2$ . Problém stochastické rekonstrukce je zadán tak, že máme k dispozici jednu konfiguraci bodů v pozorovacím okně a chceme vygenerovat sadu konfigurací té výchozí co nejvíce podobné z geometrického hlediska. Tento problém lze řešit pomocí optimalizačního algoritmu, který je v práci představen. Původní verze tohoto algoritmu funguje uspokojivě pro bodové vzorky, které odpovídají realizacím stacionárních bodových procesů. Rozšíření, navržené v literatuře, které by mělo fungovat i pro případ nestacionárních vzorků, už tak uspokojivě nefunguje. Hlavním úkolem práce bylo prozkoumat, proč tomu tak je, a navrhnout možná vylepšení. Toto bylo provedeno v kapitole 3 pomocí rozsáhlé simulační studie. V kapitole čtvrté byl navíc podrobněji prozkoumán alternativní návrh z literatury – použít místo "improvement only" optimalizačního algoritmu raději MCMC simulaci z vhodného modelu bodového procesu určeného pomocí výchozího vzorku.

Práce je psána přehledně, jasně a srozumitelně, s prakticky žádnými chybami v matematice. Pro její napsání musela autorka zvládnout poměrně široký rozsah materiálu a mimo přesného a formálně korektního představení velmi oceňuji množství poznámek, vysvětlení a souvislostí, které prokazují, že nad věcmi přemýšlí a zároveň má nad nimi jistý nadhled. Simulační studie ve třetí kapitole je dobře navržena a dobře provedena, a rovněž dobře zhodnocena. Ve čtvrté kapitole byly odvozeny teoretické vlastnosti MCMC algoritmu použitého pro stochastickou rekonstrukci a rovněž byl simulačně prozkoumán. Zadání práce bylo jistě splněno a jediná skutečnost, která systematicky kazí dojem je nezanedbatelné množství překlepů. I přes tuto výtku je ale možno prohlásit práci za velmi zdařilou.

Práce zcela jistě splňuje všechny požadavky kladené na diplomovou práci. Proto ji doporučuji jako takovou uznat.

Některé podrobnější připomínky, resp. dotazy k práci následují níže.

- Mám trochu problém s motivací, kdy je jako možná aplikace algoritmu stochastické rekonstrukce uváděno jeho možné použití pro simulation-based testy. Zcela jistě není možné použít výsledné bodové vzorky pro simulační testy hypotézy, kterou není možné dostatečně specifikovat, nebo pro průzkum sampling variation odhadovaných statistik pro model, ze kterého pochází výchozí vzorek. Je možné je použít pro podmíněnou simulaci, ale nikoli pro statistické procedury. Ten příklad ze stránky 6 není dobrý.
- Líbí se mi velmi provedení simulačních testů ze sekce 3.2. To je zcela správně. Jen by mne zajímalo, proč je volen tak malý počet vygenerovaných vzorků z modelu, které jsou použity za základ pro stochastickou rekonstrukci. Při tak malém počtu může být vliv toho, jak vypadá ten který náhodný vzorek, poměrně velký.
- Na to navazuje poznámka ke straně 40 - pokud jsou 3 z 15 výchozích konfigurací hodny zamítnutí, není něco špatné? Nebylo by lepší mít sadu výchozích realizací z modelu, kde by počet „extrémních“ konfigurací lépe odpovídal jejich průměrnému počtu  $\leq \alpha$ ? Takto „správně“ stochasticky zrekonstruované konfigurace blízké těm výchozím extrémním mohou být „správně“ zamítnuty.

- Stochastická rekonstrukce je optimalizační procedura, nikoli statistická. Je možné a dokonce rozumné použít pro konstrukci funkcionálu energie veličinu, která je schopna zachytit žádoucí vlastnost vzorku - takže pro nehomogenitu přímo empirickou funkci intenzity. Její opětovné použití v simulačním testu zkoumající kvalitu rekonstrukce nevadí. Vy totiž neodhadujete parametry, vy se snažíte vyprodukovat bodové vzorky co nejvíc se podobající tomu výchozímu. A je chyba nevyužít charakteristiku (např.  $\hat{\lambda}$ ), která vám to umožní.
- Ten třetí model, zkoumaný v kapitole 3 – transformovaný Matérnuv hardcore model, se od těch dvou předchozích liší nejen tím, že není SOIRS, ale také tím, že je lokálně anizotropní a to v různých místech okna s různou anizotropií. Ale hlavně jeho realizace vcelku připomínají realizace z lokálně škálovaného modelu, pro který existuje dobře definovaná lokálně škálovaná  $F$ -funkce a distribuční funkce vzdálenosti k nejbližšímu sousedu. Při použití těchto modifikovaných funkcí by mohly funkcionály energie  $E_5$  a  $E_7$  fungovat dobře i zde.
- str.67 – proč volíte cílové rozdělení MCMC algoritmu homogenní, když je váš výchozí bodový vzorek nehomogenní?
- obr.4.1 je velmi zvláštní – máte nějaké vysvětlení?

RNDr. Michaela Prokešová, Ph.D.